

White Paper Report

Report ID: 109103

Application Number: HD-51668-13

Project Director: Jeffrey Rydberg-Cox (rydbergcoxj@umkc.edu)

Institution: University of Missouri, Kansas City

Reporting Period: 6/1/2013-1/31/2015

Report Due: 4/30/2015

Date Submitted: 4/29/2015

Final White Paper

Performance Summary

Our Digital Humanities Start-Up Grant *A Digital Studio for the Optical and Chemical Analysis of Early Printed Books* ran from June 1, 2013 until January 31, 2015. When we proposed this project to the NEH, it was a true start-up. The investigators had long worked together in their administrative and teaching capacities at the University of Missouri-Kansas City but had never conducted a joint research project. Funding from NEH allowed us to purchase essential lab equipment to launch the project, conduct a first round of research that allowed us to focus our investigations more precisely, obtain additional funding from within our University, and develop the beginnings of a prominent public profile for our project at UMKC, within the University of Missouri system, and among a national and international scholarly community. The results of our work are available online at <http://daedalus.umkc.edu/codices>

Project Goals and Accomplishments

Initial Project Goals

The goals for *A Digital Studio for the Optical and Chemical Analysis of Early Printed Books* were as follows:

1. Create a digital image book of a four-volume guide for priests entitled *Summa Theologica* by the Florentine archbishop Antonino Pierozzi, which was printed in Nuremburg using moveable type.
2. Conduct spot-level densitometry and spectroscopy on elements in this book.
3. Image selected pages from this book at specific frequencies in the ultraviolet-visible-near-infrared spectrum.
4. Use the resulting data and images to create a digital studio that will include interactive tutorials and demonstrations that allow humanists to explore and understand the principles of optical and chemical analysis, while also encouraging scientists to explore and understand the principles of book history and codicology.

When we proposed this project to the NEH, it was truly a start-up project. The three investigators had worked together for many years in administrative capacities at the University of Missouri-Kansas City but they had never conducted a joint research project. Funding from NEH allowed us to purchase essential lab equipment to begin this project, conduct a first round of research that allowed us to focus our investigations more precisely, obtain additional funding from within our University, and develop the beginnings of a public profile for our project at UMKC, within the University of Missouri system, and among a national and international scholarly community.

Project Accomplishments

With NEH funding we were able to successfully accomplish all of these tasks:

- 1) As part of an interactive image-book, our project web pages have visible light images of the four *Summa Theologica* volumes available:
<http://daedalus.umkc.edu/codices>.
- 2) We also started to develop the capacity to analyze select elements from the books using densitometry and, more generally, reflectance spectroscopy. The bulk of this work focused on developing non-invasive techniques and tools to position a fiber optic sensor on a specific area. It lead to the development of a robotic apparatus that can precisely move the sensors across an XY axis allowing for better analysis of larger areas. Currently, while we have the capability to perform spectroscopy at a spot manually, the ability to scan a particular area using the XY robot with a consistent orientation and distance to the papers is still in development.

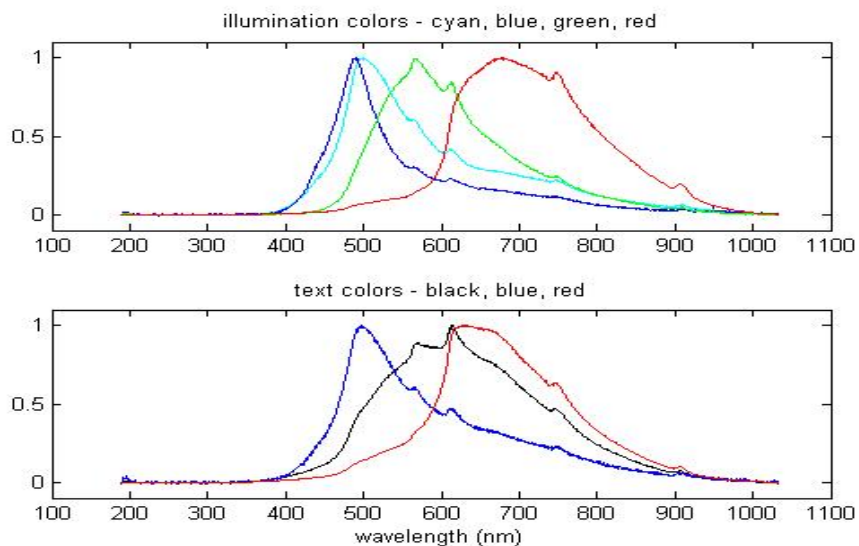


Figure 1: Reflectance spectra of text and illumination colors. Note the text and illumination reds and blues are spectrophotically distinct from each other clearly indicating that different inks were used in the text vs. illumination.

- 3) We extensively examined select pages from this book under specific frequencies of light. As we began the grant project we had a broad list of potential areas of

investigation, such as pages that were used more heavily than others, preparatory drawings underneath the illuminations in the book, details about the relative age of the marginalia or the locations where the ink used for marginalia might have been sourced. As we worked on this book, our interests coalesced around two topics: 1) principal component analysis of the multi-spectral images and 2) a more focused set of questions related to the identification of watermarks.

3a) PCA: In traditional color imaging, each image pixel is associated with intensity values from three separate bands in the visible spectrum: Red (600 to 700 nm), Green (500 to 600 nm), and Blue (400 to 500 nm). In multispectral and hyperspectral imaging, images are captured with a higher spectral resolution using narrow spectral bands within a broader spectral range. Our goal is to detect and discriminate features of the image using the integration of reflected photons at multiple wavelengths. The imaging process generates a multidimensional dataset which is traditionally referred to as a HyperCube where the third dimension is the number of spectral bands. One of the most commonly used methods for dimensionality reduction is principal component analysis (PCA). Mathematically, it is based on an eigen decomposition of the correlation matrix of the data. The principal components produced are ordered by eigenvalue (from most significant to least significant).

Our initial exploratory work was performed on hyperspectral data acquired with a commercial DSLR Canon camera equipped with an electronic liquid crystal display filter. The images were captured in a laboratory environment using a broadband light source. The use of PCA is demonstrated on a HyperCube of images of a book page

containing a multi-color illumination. The PCA algorithm was implemented in Matlab (The MathWorks Inc.).

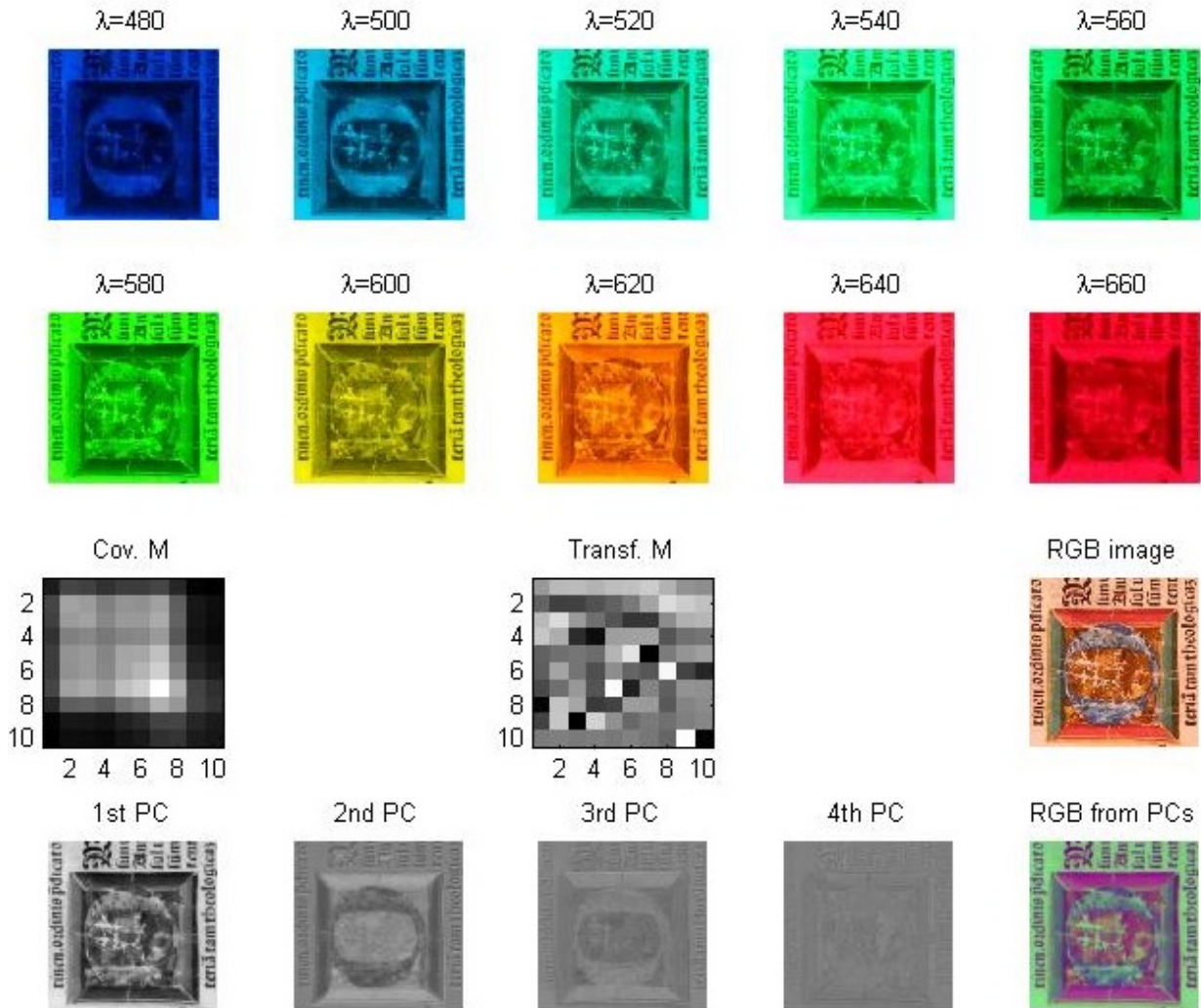


Figure 2: Dimensionality reduction with PCA. From left to right the top two rows present the original hyperspectral images acquired in 20 nm steps in the range 480 to 660 nm of the visible spectrum. The third row presents the calculated covariance and transform matrices by the principal component analysis algorithm. On the right end of the third row is the original RGB image. On the fourth row from left to right are presented the four top ranking principal components (PCs). As expected, the first component contains the most information, but some additional information can be gained from the higher principal components. The last image on the fourth row is the composite RGB image in which the first PC is mapped to the green channel, the second PC to the red, and the third PC to the blue channel.

3b) Watermarks: This work evolved in phases: first we realized that human researchers could identify watermarks much more easily under several specific

frequencies of light. As a result, we began to focus more intensively on watermark identification with two Humanities graduate students developing a catalog of all of the watermarks in our book, with attention to using illumination at different frequencies to identify watermarks. This exercise, in turn, caught the interest of a Computing and Engineering graduate student who developed algorithms for the automatic extraction and identification of watermarks from the images captured.

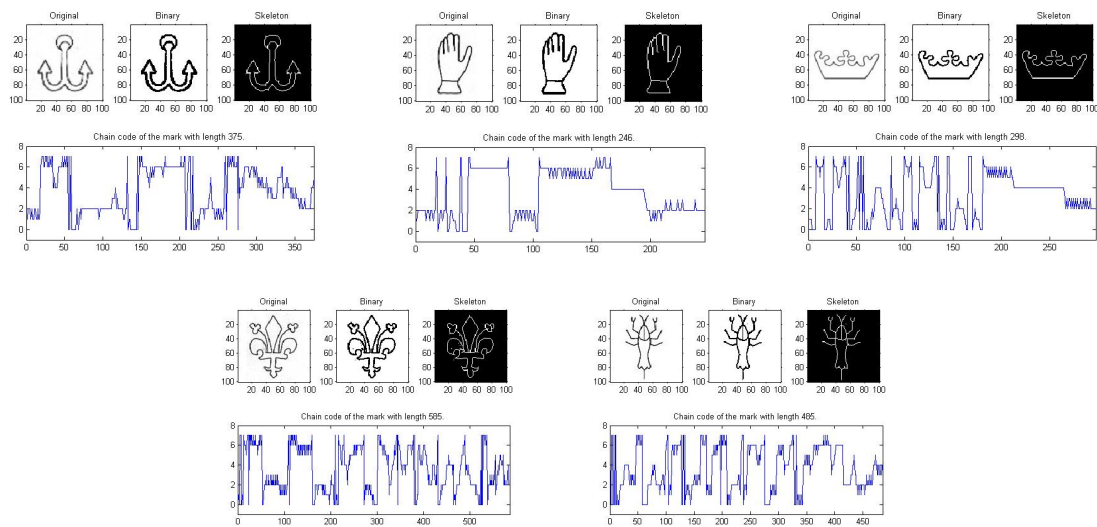


Figure 3: Numerical chain code describing various watermarks. Chain code is used to describe the content of the watermark image much more compactly and may be easily stored in a searchable database that is independent of image properties like pixel resolution etc.

- 4) In addition to the technical analysis of the book, our team also completed four substantial tools and demonstrations that are designed to introduce non-humanists to the principles of book history and non-scientists to the principles of optical and chemical analysis. The first of these demonstrations was a complete history of the *Summa Theologica* that included sections on the book's provenance and its codicological and paleographical details. This book history is approximately ten thousand words and is available as part of our web pages. We

created a companion document that describes the process of writing a book history such as the one we have created for the *Summa Theologica*. This document runs some ten thousand words and could be used in book history courses as an introduction and practical guide to writing a book history. Based on our investigations using specific light frequencies, we have also created a document containing detailed descriptions of the watermarks and paper in this book. Finally, we have developed an interactive tutorial that illustrates some basic principles of optical analysis and how it can be applied to early printed books.

Publicity and Audiences

In the first phases of our project, we have reached regional and international audiences that consist of both academics and a broader general public. Regionally, we have presented our work in progress at two meetings of the Mid-America Medieval Association in 2014 and 2015. Our presentation in 2014 was a narrative overview of our project while the 2015 presentation was a hands-on demonstration of our lab space. We reached an international audience when two of our graduate students presented at a conference in London celebrating the 50th Anniversary of the Printing Historical Society, *Landmarks in Printing: From Origins to the Digital Age*. Their presentation, entitled “Reimagining Incunables: How Accessible Digital Technology Illuminates the Production Process of Antoninus’ *Summa Theologica*,” covered their work on the history of the four-volume *Summa Theologica* and the identification and cataloging of watermarks. Our proposal to present our work at the July 2015 Early Book Society meeting in Oxford, UK, has also been accepted.

Our work has also reached a more general audience with the support of the research services office at UMKC. Our group presented at a research showcase designed to highlight areas of research at UMKC and our presentation was featured on the main UMKC home page in December 2014. In addition, our research office publishes a glossy magazine which is mailed to UMKC supporters and members of the Kansas City community. Our work is the subject of a feature article that will appear in the issue of this magazine that will be published in April of 2015.

The project web site, which remains the primary vehicle that we will use to reach a large audience, is also one of the deliverables of this grant. The web pages were only made publically available in the first part of 2015, but we hope that these will begin to reach a broad audience as our work continues.

Evaluation

We have not yet engaged in a formal evaluation of the project per se. The peer review panels for the two rounds of internal funding have served as a form of proxy evaluation that suggests that our work is on the right track. We hope to engage in a more formal round of evaluation as we move forward with the project.

Continuation of the Project

The project team plans to remain intact and continue working on this project in the future. The internal funding from the University of Missouri system runs through August of 2015. With this, and additional support from another University of Missouri granting agency—the Interdisciplinary Intercampus Research Grant—and with the collaboration with Stuart Hinds, Director of LaBudde Special Collections at UMKC, our team is turning its attention to three new books, the first a Gregorian

Chant Book held in the University of Missouri-Kansas City libraries that combines both manuscript and print materials from the 15^h to the 17th centuries. The book contains liturgical music that would have been used for the daily offices, and other services. It provides an excellent test case for multispectral analysis techniques because it allows us to investigate the multiple scribal hands that appear in the book; significant amount of marginal annotation, and most exciting, the corrections, additions, and deletions from the manuscript pages where a scribe has scraped away or painted over an earlier version of the text and replaced it with a new version. As part of the UM System support, we have also begun collaborating with Alla Barabtarlo, the Head of Special Collections and Rare Books, Ellis Library, University of Missouri-Columbia, Julie Christensen, a PhD student in the English department at the University of Missouri-Columbia, and Anne Stanton, a Professor of Art History at the University of Missouri-Columbia campus who are using our methods and who will develop a similar lab to study a 1501 missal from Prague and a 1624 Menaion, or Russian service book.

Long Term Impact

Our team is in the process of generating a specific spin-off platform that will incorporate our work into a university-wide program in which undergraduate students engage in research alongside faculty. We are also creating a 200-level undergraduate course proposal that incorporates all of the elements from our grant project, book history, principles of optical and chemical analysis, and book history. Students who take this course will be asked to engage in a research project based on a manuscript or early printed book and those who complete the course will be eligible

to apply for stipends that will allow them to work on future book projects as part of our team.

Grant Products

Project Web Site: <http://daedalus.umkc.edu/CODICES>

Project White Papers:

History of the *Summa Theologica*: online at:

<http://daedalus.umkc.edu/codices/papers/SummaTheologicaBookHistory.CodicesPaper1.2015.pdf>

Description of how to construct a book history online at:

<http://daedalus.umkc.edu/codices/papers/HowToWriteAHistoryoftheBook.CodicesPaper2.2015.pdf>

Summa Theologica Book Images:

http://daedalus.umkc.edu/codices/SummaTheologica/summa_v1.html